Cross-Gramian Based Combined State and Parameter Reduction

Christian Himpe* Mario Ohlberger*

Abstract

An accepted model reduction technique is balanced truncation, by which negligible states of a linear system of ODEs are determined by balancing the systems controllability and observability gramian matrices. To be applicable for nonlinear system this method was enhanced through the empirical gramians, while the cross gramian conjoined both gramians into one gramian matrix. This work introduces the empirical cross gramian for square Multiple-Input-Multiple-Output systems as well as the (empirical) joint gramian. Based on the cross gramian, the joint gramian determines, in addition to the Hankel singular values, the parameter identifiability allowing a combined model reduction, concurrently reducing state and parameter spaces. Furthermore, a controllability and an observability based combined reduction method are presented and the usage of empirical gramians is extended to parameter reduction in (Bayesian) inverse problems. All methods presented are evaluated by numerical experiments.

Keywords: Nonlinear, Control, System, Parameter, Model, Order, Reduction, Balanced, Truncation, Empirical, Cross, Joint, Gramian, Grammian, MIMO, Sensitivity, Identifiability, Controllability, Observability, PCA, POD, ODE

MSC: 93B11, 93B30, 93C10

ACM: G.1.3

1 Introduction

Efficient reduction of large-scale nonlinear control systems is a challenging task. Even more in the case of parametrized systems, in an inverse problem setting, where a combined reduction of parameters and states is targeted. For instance, a network with many nodes and unknown connectivity, modeled as a parametrized nonlinear dynamic system, requires much time during parameter estimation due to system size and parameter count. With the aim to lower computational complexity, the parameter and state spaces are confined

*Contact: `christian.himpe@wwu.de`, `mario.ohlberger@uni-muenster.de`, Institute for Computational and Applied Mathematics at the University of Münster, Einsteinstrasse 62, D-48149 Münster, Germany

to low-dimensional subspaces. One class of model order reduction techniques is concerned with determining projections to such subspaces, mapping the high-dimensional model to a reduced surrogate model. Here, the projection based method of balanced truncation is the origin of the proposed combined reduction of state and parameter space.

Since the number of a systems inputs and outputs usually remains fixed, the maps to and from the intermediary states characterize the reducibility of a system ([1]). The balanced truncation approach, introduced in [2], categorizes the states of a system in terms of controllability and observability. From the input-to-state map the systems controllability is computed, quantifying how well a state is driven by input. The state-to-output map is used to compute the systems observability, quantifying how well changes in a state are reflected by the output. Excluding the least controllable and observable states from a model motivates a projection from the full to a truncated model. Given a linear time-invariant control system,

$$\begin{aligned}\dot{x}(t) &= Ax(t) + Bu(t)\\ y(t) &= Cx(t),\end{aligned}$$

this systems controllability and observability can be assessed through the associated gramian matrices W_C and W_O, which are classically computed as the solutions of the Lyapunov equations $AW_C + W_C A^* = -BB^*$ and $AW_O + W_O A^* = -C^*C$. To make a compound statement about controllability and observability, W_C and W_O have to be balanced. The singular values of the resulting balanced gramian correspond to the Hankel singular values of the system, with their magnitude describing how controllable and observable the associated state is. A description of the balancing procedure can be found, for example, in [3] or [4].

This work focuses on cross gramian based methods for model reduction. The cross gramian was introduced in [5]; it combines controllability and observability information into a single gramian matrix and is elaborately described in [6]. The empirical gramians were introduced in the works [7], [8] and enable the computation of gramian matrices for nonlinear systems by mere basic matrix operations. This concept was extended among others, in [9] allowing nonlinearity measurement and in [10] providing more general input signals. Particularly noted should be [11] and [12] for developing the empirical cross gramian for SISO systems in the context of sensitivity analysis. For the gramian based parameter reduction, the groundwork has been laid by [13], from the observability standpoint and by [14], from the controllability point of view.

To begin, the cross gramian and its properties are reviewed in section 2. Then the empirical cross gramian (for MIMO systems) is developed in section 3. Section 4 introduces combined state and parameter reduction in two variants. First, a controllability and second, an observability based approach; the latter is enhanced to a cross gramian based combined reduction, that is presented in section 5. A procedure to reduce parameters spaces in Bayesian inverse problems, prior to parameter estimation, is suggested in section 6. Finally, numerical experiments are conducted in section 7 comparing the here presented methods for linear and nonlinear systems.

2 Cross Gramian

The cross gramian W_X (also known as: W_{CO}) was introduced in a sequence of works ([5], [15], [16], [17], [18], [19], [20]) and encodes controllability and observability into a single matrix, yet it can only be computed for square systems, meaning a system with the same number of inputs and outputs. In this case the cross gramian is the smallest semi-positive solution of the Sylvester equation $AW_X + W_X A = -BC$, of which approximate solutions were discussed in [21], [22] and [23]. The cross gramian can also be specified by the product of input-to-state and state-to-output maps:

$$W_X = \int_0^\infty e^{A\tau} BCe^{A\tau} d\tau, \tag{1}$$

in case the underlying system is asymptotically stable. If the system is further on symmetric, the following relation between the cross, the controllability and observability gramian holds ([5]):

$$W_X^2 = W_C W_O \Rightarrow |\lambda(W_X)| = \sqrt{\lambda(W_C W_O)}.$$

And as shown in [22, Lemma 2.1], the eigenvalues of the cross gramian correspond to the Hankel singular values.
While a for Single-Input-Single-Output (SISO) system this is always true ([5]), a linear Multiple-Input-Multiple-Output (MIMO) system not only requires the same number of inputs and outputs, but also the system gain $G = -CA^{-1}B$ to be symmetric ([21]), which is equivalent to the existence of a symmetric transformation J, with $AJ = JA^T$ and $B = C^T J$. Trivially for $J = \mathbb{1}$ the system would be restricted by $A = A^T$ and $B = C^T$.
As presented in [5] the trace of the cross gramian equals half the gain of a SISO system

$$\mathrm{tr}(W_X) = -\frac{1}{2} CA^{-1}B.$$

The trace being the sum of eigenvalues, this relationship associates the eigenvalues with the system gain. This result was used in [11] and [12] for parameter identification purposes, using the system gain as a sensitivity measure. For MIMO systems, here an extension to [5, Theorem 3] is developed:

Corollary
Given a linear, square MIMO system, the trace of the cross gramian relates to the system gain as follows:

$$\mathrm{tr}(W_X) = -\frac{1}{2}\,\mathrm{tr}(CA^{-1}B).$$

Proof.
For an asymptotically stable system, the trace of the cross gramian (in the form of equation 1) is given by:

$$\mathrm{tr}(W_X) = \mathrm{tr}(\int_0^\infty e^{A\tau} BCe^{A\tau} d\tau)$$

$$= \int_0^\infty \mathrm{tr}(e^{A\tau} B C e^{A\tau}) d\tau,$$

and since the trace of a product of compatible matrices is commutative,

$$\begin{aligned}
&= \int_0^\infty \mathrm{tr}(C e^{A\tau} e^{A\tau} B) d\tau \\
&= \mathrm{tr}(\int_0^\infty C e^{2A\tau} B d\tau) \\
&= \mathrm{tr}(C \int_0^\infty e^{2A\tau} d\tau B). \\
&= \mathrm{tr}(C(-\frac{1}{2} A^{-1}) B) \\
&= -\frac{1}{2} \mathrm{tr}(C A^{-1} B).
\end{aligned}$$

□

Employing the cross instead of controllability and observability gramian, means only a single gramian has to be computed. And since no balancing is required, the truncation procedure can be simplified to direct truncation ([24], [25, Ch. 7], [6, Ch. 12.3]). The projection is given by the singular value decomposition of the cross gramian,

$$UDV = \mathrm{svd}(W_X).$$

Then U is truncated based upon the singular values D, yielding a projection matrix $\hat{U}$. Applying $\hat{U}$ to A, B, C results in the reduced system:

$$\Rightarrow \begin{cases} \dot{x}_r(t) & = \hat{U}^T A \hat{U} x_r(t) + \hat{U}^T B u(t) \\ y(t) & = C \hat{U} x_r(t). \end{cases}$$

Lastly, the cross gramian has the benefit of conveying more information then the controllability and observability gramian ([16]).

3 Empirical Cross Gramian

For nonlinear models,

$$\begin{aligned}
\dot{x} &= f(x, u, \theta) \\
y &= g(x, \theta) \\
x(0) &= x_0,
\end{aligned}$$

the above procedure is not viable. In [7], [8] and [26] the concept of empirical (controllability and observability) gramians was introduced. This is an extension of the PCA method which uses the SVD of a snapshot matrix ([27]). Subsequently this approach and its field of application was advanced by [28], [29], [30] and [31]. The empirical gramians not only correspond to the classic gramians for linear systems ([7]), but might even carry more detailed information on the underlying system, as shown in [32]. Empirical gramians are based on averaging

the output of a system that is perturbed in inputs and initial states. Initially only for delta impulse type input, this ansatz was extended in [10] for more general input configurations under the name of controllability and observability covariance matrices.
Next, the necessary perturbation sets are defined; these should reflect the operating range of the underlying system. E and F are sets of standard directions for the inputs and initial states. Sets S and T are orthogonal transformations (rotations) to these standard directions of inputs and initial states respectively, while Q and R hold scales to these directions:

$$
\begin{aligned}
E &= \{e_i \in \mathbb{R}^m; \|e_i\| = 1; e_i e_j = 0; i = 1, \dots, m\}, \\
F &= \{f_i \in \mathbb{R}^n; \|f_i\| = 1; f_i f_j = 0; i = 1, \dots, n\}, \\
S &= \{S_i \in \mathbb{R}^{m \times m}; S_i^* S_i = \mathbb{1}; i = 1, \dots, s\}, \\
T &= \{T_i \in \mathbb{R}^{n \times n}; T_i^* T_i = \mathbb{1}; i = 1, \dots, r\}, \\
Q &= \{q_i \in \mathbb{R}; q_i > 0; i = 1, \dots, q\}, \\
R &= \{r_i \in \mathbb{R}; r_i > 0; i = 1, \dots, r\}.
\end{aligned}
$$

The empirical cross gramian for SISO systems was introduced in [11]; here, as a new contribution, this concept will be generalized to square MIMO systems. Hence the scope of the cross gramian is extended to nonlinear systems. For square MIMO systems the empirical cross gramian is then defined by:

Definition (Empirical Cross Gramian)
For sets E, F, S, T, Q, R and the steady state $\bar{x}, \bar{u}(\bar{x}), \bar{y}(\bar{x})$ the **empirical cross gramian** relating to the output y^{kli}, the states x^{mhj} and the input $u(t) = c_m S_h f_j \delta(t) + \bar{u}$ as well as the initial state $x(0) = d_k T_l e_i + \bar{x}$ is given by:

$$
\begin{aligned}
\hat{W}_X &= \frac{1}{mrsv} \sum_{h=1}^{m} \sum_{k=1}^{v} \sum_{l=1}^{r} \sum_{m=1}^{s} \frac{1}{d_k c_m} \int_0^\infty \Psi^{hklm}(t) dt \\
\Psi_{ij}^{hklm} &= e_i^T T_l^T \Delta x^{mhj}(t) f_j^T S_h^T \Delta y^{kli}(t) \\
\Delta x^{mhj}(t) &= (x^{mhj}(t) - \bar{x}^{mhj}) \\
\Delta y^{kli}(t) &= (y^{kli}(t) - \bar{y}^{kli}).
\end{aligned}
$$

Along the lines of [7] and [11], the equality of the cross gramian and the empirical cross gramian of a linear MIMO system is shown next:

Lemma (Empirical Cross Gramian)
For any nonempty sets S, T, Q, R the empirical cross gramian $\hat{W}_X$ of an asymptotically stable linear control system is equal to the classic cross gramian.

Proof.
For an asymptotically stable linear control system, $\Delta x(t) = x(t) = e^{At} B u(t)$

and similarly $\Delta y(t) = y(t) = Ce^{At}x_0$, thus:

$$\begin{aligned}
\Psi_{ij}^{hklm} &= e_i^T T_l^T (e^{At} B c_m S_h f_j) f_j^T S_h^T (C e^{At} d_k T_l e_j) \\
&= d_k c_m e_i^T T_l^T e^{At} B C e^{At} T_l e_i \\
\Rightarrow \Psi^{hklk} &= d_k c_m T_l^T e^{At} B C e^{At} T_l \\
\Rightarrow \hat{W}_X &= \frac{1}{v} \sum_{k=1}^{v} \int_0^\infty e^{At} B C e^{At} dt \\
&= W_X.
\end{aligned}$$

□

Following [9], a discrete representation of the empirical cross gramian is given here too, due to the situation that data is collected at discrete times.

Definition (Discrete Empirical Cross Gramian)
For sets E,F,S,T,Q,R the **discrete empirical cross gramian** relating to the output y^{kli}, the states x^{hjl} and the input $u(t) = c_h S_j e_l \delta(t) + \bar{u}$ as well as the initial state $x(0) = d_k T_l e_i + \bar{x}$ is given by:

$$\begin{aligned}
\tilde{W}_X &= \frac{1}{mrsv} \sum_{h=1}^{m} \sum_{k=1}^{v} \sum_{l=1}^{r} \sum_{m=1}^{s} \frac{\Delta t}{d_k c_m} \sum_{g=0}^{G} \Psi^{hklm} \\
\Psi_{ij}^{hklm} &= e_i^T T_l^T \Delta x_g^{mhj} S_h^T f_j^T \Delta y_g^{kli} \\
\Delta x_g^{mhj} &= (x_g^{mhj} - \bar{x}^{mhj}) \\
\Delta y_g^{kli} &= (y_g^{kli} - \bar{y}^{kli}).
\end{aligned}$$

The computational complexity depends largely on the number of scales and rotations of perturbations as well as the cost of integration.

4 Combined State and Parameter Reduction

In this section two methods for combined model reduction, allowing simultaneous reduction of state and parameter spaces are introduced; an observability based and a controllability based ansatz.
In [13] the identifiability gramian was introduced, which extends the concept of observability from states to parameters. Augmenting the states of a given system by including its parameters as constant states, the parameters are treated like states:

$$\begin{aligned}
\dot{\breve{x}} &= \begin{pmatrix} \dot{x} \\ \dot{\theta} \end{pmatrix} = \begin{pmatrix} f(x,u,\theta) \\ 0 \end{pmatrix} \\
y &= g(x,\theta) \\
\breve{x}(0) &= \begin{pmatrix} x(0) \\ \theta \end{pmatrix}.
\end{aligned}$$

These augmented states are initialized by their parameter value. The identifiability of the parameters is obtained through the observability of these augmented

parameter-states. From the observability gramian of this augmented system, the identifiability gramian W_I can be extracted via the Schur-complement,

$$\tilde{W}_I = \left(\begin{array}{c|c} W_O & W_a^* \\ \hline W_a & W_p \end{array}\right)$$
$$\Rightarrow W_I = W_p - W_a^* W_O{}^{-1} W_a.$$

The identifiability is then given as the observability of the parameters. A parameter reducing projection is computed, similarly to the cross gramian approach, through a singular value decomposition of the identifiability gramian. Observability of the states is encoded in this combined observability gramian, too; it can be extracted as the upper-left $N \times N$ matrix from $\tilde{W}_I$. Thus after a identifiability gramian based parameter reduction, and computation of the systems controllability gramian, a state reducing balanced truncation can be applied to achieve combined reduction.
In [14] the per parameter controllability gramians were introduced for sensitivity analysis purposes, in which the parameters are treated as additional inputs. Controllability gramians are computed for the system without any parameter inputs ($W_{C,0}$), and for each parameter as sole input to system excluding other parameters as well as regular inputs ($W_{C,i}$). Then a controllability gramian of this system is given by the sum of all sub controllability gramians:

$$W_C = \sum_{i=0}^{P} W_{C,i}.$$

The sensitivity related controllability information (the traces of the associated sub-gramian $W_{C,i>0}$ of each parameter) can then be accumulated in a diagonal matrix, which is called the **sensitivity gramian** (see [33]):

$$W_S = \begin{pmatrix} \mathrm{tr}(W_{C,1}) & & 0 \\ & \ddots & \\ 0 & & \mathrm{tr}(W_{C,P}) \end{pmatrix}.$$

The singular value decomposition of W_S induces a parameter projection. In combination with an observability gramian, a balanced truncation of states subsequent to a sensitivity gramian based parameter reduction results in a combined reduction.

5 Joint Gramian

Now, aggregating the computation of controllability and observability not only for states as proposed above, but also for identifiability of parameters leads to a reduction method requiring a new single gramian. Additionally to extending the states with as many (constant) states as there are parameters; inspired by [34], the inputs are extended to include controllability of the parameters. To be able to apply the cross gramian to this extra augmented system, the outputs have to be augmented as well, to keep the system square. The augmented inputs act untransformed upon the parameters; consequently, to ensure the systems symmetry, the output transformation of the augmented outputs is represented by an identity function. This leads to the following new gramian matrix:

Definition (Joint Gramian)
An extra augmented system, with $v \equiv 0$, is given by

$$\begin{aligned}
\breve{u} &= \begin{pmatrix} u \\ v \end{pmatrix} \\
\dot{\breve{x}} &= \begin{pmatrix} \dot{x} \\ \dot{\theta} \end{pmatrix} = \begin{pmatrix} f(x,u,\theta) \\ v \end{pmatrix} \\
\breve{y} &= \begin{pmatrix} g(x,\theta) \\ \mathbb{1} \end{pmatrix} \\
\breve{x}(0) &= \begin{pmatrix} x(0) \\ \theta \end{pmatrix}.
\end{aligned}$$

The cross gramian of such extra augmented system is called the **joint gramian**.

Applying the empirical cross gramian to this setting allows to compute the empirical joint gramian, which is composed of:

$$W_J = \left(\begin{array}{c|c} W_X & W_a^* \\ \hline W_a & W_p \end{array}\right),$$

W_X being the cross gramian and W_p the parameter-related gramian. The identifiability information on the parameters can be extracted by the Schur complement:

$$W_{\ddot{I}} = W_p - W_a^* {W_X}^{-1} W_a,$$

with $W_{\ddot{I}} \in \mathbb{R}^{p \times p}$ the **cross identifiability gramian**, which is usually not equal to the identifiability gramian presented in [13], since here not only observability, but also controllability is part of the identifiability. As for W_I and W_S the associated parameter projection is computed by the singular value decomposition of $W_{\ddot{I}}$. Since the (empirical) gramians are symmetric, a variant of the Cholesky decomposition, based on [35], can be used to compute the Schur-complement:

$$\begin{aligned}
\tilde{W}_J &= W_J + \|W_J\|_1 \mathbb{1} \\
L^* L &= \mathrm{chol}(\tilde{W}_J) = \mathrm{chol} \begin{pmatrix} W_X & W_a^* \\ W_a & W_p \end{pmatrix} \\
L &= \begin{pmatrix} L_a & L_b \\ 0 & L_c \end{pmatrix} \\
\Rightarrow \tilde{W}_{\ddot{I}} &= L_c L_c^* \\
\Rightarrow W_{\ddot{I}} &\approx \tilde{W}_{\ddot{I}} - \|W_J\|_1 \mathbb{1}.
\end{aligned}$$

For a combined reduction of states and parameter no further gramian has to be computed and no balancing is required, since the cross gramian W_X for state reduction is given by the upper-left $N \times N$ matrix of the joint gramian W_J.

6 Bayesian Inversion

The empirical gramian method is especially well suited for parameter identifiability, for example see [36]. Parameter identifiability or reduction for inverse

problems is a complicated setting, since the parameter reduction should occur before their estimation. Especially for Bayesian inverse problems, empirical gramians can be utilized with ease due to prior distribution on the unknown parameters, which is presented next: Assuming a Gaussian setting, the provided prior knowledge, given by mean and covariance can be employed to define the configuration of the parameters steady states and perturbations. Using the prior mean, the steady states of the augmented parameter states are set,

$$\begin{aligned}\dot{x} &= f(x, u, \mathrm{E}[p(\theta)]) \\ y &= g(x, \mathrm{E}[p(\theta)]),\end{aligned}$$

while the prior covariance defines the perturbation scales, given a sequence $a_i \in [0, 1]$:

$$\begin{aligned}Q_{\theta_j} &= \{q_i = a_i \sqrt{\mathrm{Var}[p(\theta_j)]}; i = 1 \dots m\}, \\ R_{\theta_j} &= \{r_i = a_i \sqrt{\mathrm{Var}[p(\theta_j)]}; i = 1 \dots n\}.\end{aligned}$$

For the actual parameter reduction, after the identifiability has been determined, two reduction strategies are available. First, following the truncation procedure, by projecting the priors to a low-dimensional subspace and estimating only a low number of parameters. Second, by identifiability analysis, using the eigenvectors associated to the eigenvalues of the gramian, the reducible parameters are excluded from the estimation procedure and kept at their prior mean value. As described above not only parameter reduction, but also combined reduction of states and parameters can be applied in this pre-estimation Bayesian inverse problem setting.

7 Implementation and Numerical Experiments

For an efficient implementation the structure of the gramian computation is exploited. First, the empirical gramians allow extensive parallelization. Each combination of direction, transformation and scale can be processed separately yielding a sub-gramian. Second, the assembly of each sub-gramian can be comprehensively vectorized, since it consists of vector additions, inner- and outer-products. The output gramian is the normalized accumulation of all sub-gramians. For further details about the implementation refer to [33].

The empirical gramian framework can be found at `http://gramian.de/emgr.m` and is compatible with **Octave** and **Matlab** while the source code, used for the following experiments, can be found at `http://gramian.de/himpe13a.tgz`.

The empirical cross gramian and empirical joint gramian are verified with a (parametrized) linear system,

$$\begin{aligned}\dot{x}(t) &= A_\theta x(t) + Bu(t) \\ y(t) &= Cx(t),\end{aligned}$$

and tested with a (parametrized) nonlinear system, with the nonlinearity being a component-wise applied sigmoid function, here a hyperbolic tangent (see [37]);

$$\begin{aligned}\dot{x}(t) &= A_\theta \tanh(x(t)) + Bu(t) \\ y(t) &= Cx(t).\end{aligned}$$

In both cases the matrix A_θ is fully parametrized, meaning each component of A_θ is a separate parameter, hence $\theta \in \mathbb{R}^P = \mathbb{R}^{N^2}$. Here the linear and nonlinear system are chosen to have $I = O = 4$ inputs and outputs as well as $N = 32$ states and thus $P = N^2 = 1024$ parameters. Regardless of the Hankel singular values, state reductions reduce to $N_R = 4$ states and parameter reductions to $P_R = 16$ parameters. The parameters constituting A_θ are generated randomly, but stability and symmetry of the system are ensured. For parameter and combined reductions the experiments are treated as Bayesian inversion, reducing the parameter space before estimation that is initialized with the prior mean. Input matrix B is also random and the output matrix is given by $C = B^T$ to guarantee symmetry. Following, each figure (fig. 1, fig. 2, fig. 3, fig. 4, fig. 5, fig. 6) shows the relative error of the reduced to the full model for each state over time for an exemplary system.
The state reduction using the empirical cross gramian (W_X) is compared to balanced truncation of empirical controllability gramian (W_C) and empirical observability gramian (W_O), and also to a variant of balanced proper orthogonal decomposition from [38] and [39].

	offline time $[s]$	online time $[s]$	$L2$-error
original system	0	0.0061	0
balanced POD	0.1279	0.0112	0.0139
balanced truncation (W_C, W_O)	0.9747	0.0110	0.0223
direct truncation (W_X)	0.6021	0.0111	0.0149

Table 1: ***Linear*** *State Reduction, averaged over 10 systems.*

	offline time $[s]$	online time $[s]$	$L2$-error
original system	0	0.0063	0
balanced POD	0.1330	0.0125	0.0154
balanced truncation (W_C, W_O)	1.0462	0.0117	0.0370
direct truncation (W_X)	0.6538	0.0117	0.0151

Table 2: ***Nonlinear*** *State Reduction, averaged over 10 systems.*

For either, the linear (table 1) and nonlinear (table 2) state reduction the cross gramian based direct truncation requires less time than the balanced truncation using controllability and observability gramian. Naturally, balanced POD is faster, but produces a lower error only for the linear case; for the nonlinear state reduction the empirical cross gramian reduces with the least error.
For parameter and combined reduction the parameters constituting A_θ are treated as unknown and are initialized with flat priors. Thus each online phase solves an inverse problem by estimating the parameters using unconstrained optimization. The parameter reduction performance, of the empirical joint gramian (W_J) is assessed in comparison to the empirical sensitivity gramian (W_S, see [33]), and the empirical identifiability gramian (W_I, [13]).
In all cases (see table 3 and table 4), the online time is reduced by approximately two orders of magnitude. Linear and nonlinear parameter reduction is completed

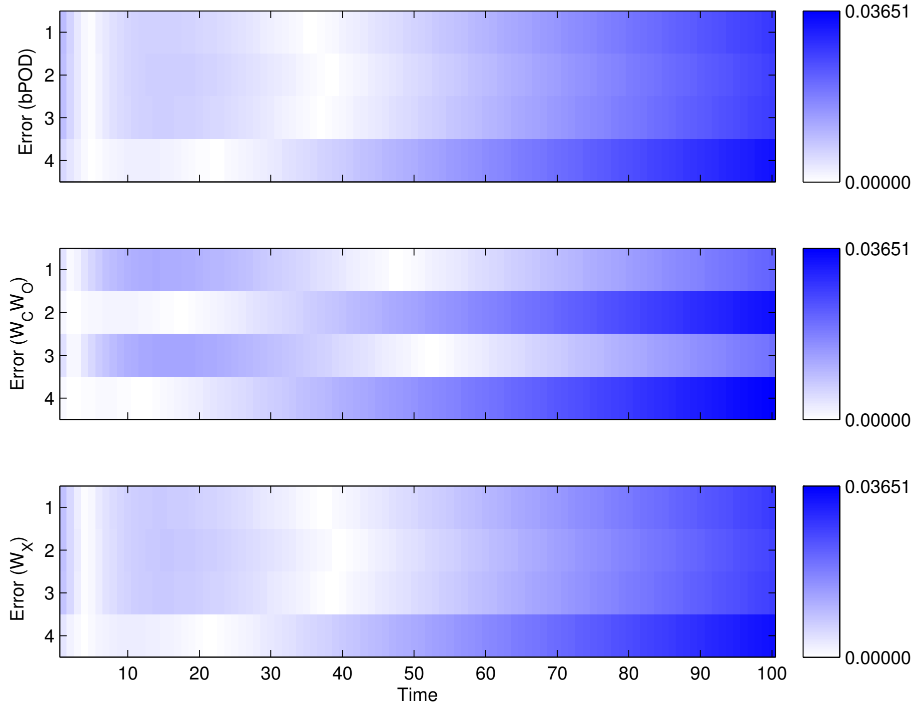

Figure 1: *Exemplary absolute error comparison of **state reduction** for a **linear** system by balanced POD, balanced truncation (using empirical controllability gramian and empirical observability gramian) and direct truncation (using the empirical cross gramian).*

	offline time $[s]$	online time $[s]$	$L2$-error
original system	0	353.6360	0.0228
direct truncation (W_S)	38.2590	9.3704	1.5826
direct truncation (W_I)	133.5540	5.9549	0.3361
direct truncation (W_J)	1143.8626	6.0806	0.3204

Table 3: ***Linear** Parameter Reduction, averaged over 10 systems.*

	offline time $[s]$	online time $[s]$	$L2$-error
original system	0	355.4448	0.0091
direct truncation (W_S)	40.4850	8.6330	0.9922
direct truncation (W_I)	136.2866	7.5322	0.0988
direct truncation (W_J)	1163.3324	7.5328	0.0578

Table 4: ***Nonlinear** Parameter Reduction, averaged over 10 systems.*

with the least error by the joint gramian, yet requires significantly more **offline** time than the sensitivity and identifiability gramian. The sensitivity gramian W_S demands the least offline time, but the resulting error is up to an order of

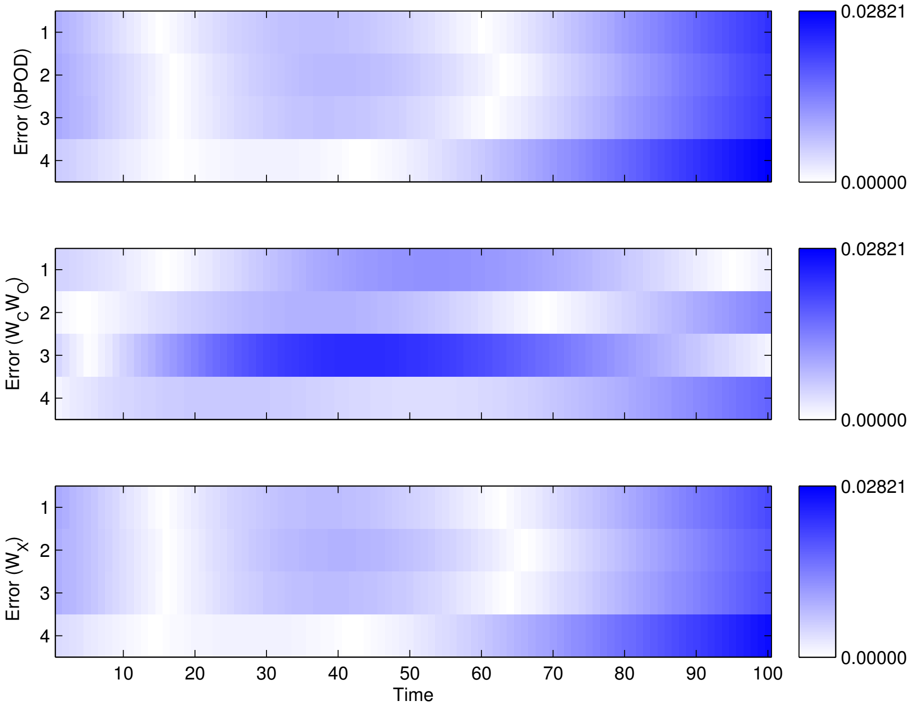

Figure 2: *Exemplary absolute error comparison of* ***state reduction*** *for a* ***non-linear*** *system by balanced POD, balanced truncation (using empirical controllability gramian and empirical observability gramian) and direct truncation (using the empirical cross gramian).*

magnitude larger compared to the identifiability gramian W_I, with its offline time still being below the original models **online** time.
Finally, the combined state and parameter reduction of the empirical joint gramian is benchmarked against the two variants of combined reduction by balanced truncation described above. The first variant reduces the parameter space by the empirical sensitivity gramian; since a controllability gramian is already computed in the process, then with an empirical observability gramian the state space is reduced by balanced truncation. The second variant reduces the parameter space by the empirical identifiability gramian, then a state space reduction, utilizing balanced truncation of the empirical controllability gramian and observability gramian, latter being a byproduct of the identifiability gramian.

	offline time $[s]$	online time $[s]$	$L2$-error
original system	0	353.6360	0.0228
balanced truncation (W_S, W_O)	39.1512	12.9656	0.0611
balanced truncation (W_C, W_I)	133.6889	6.0815	0.1117
direct truncation (W_J)	1143.8660	6.6106	0.0498

Table 5: ***Linear*** *Combined Reduction, averaged over 10 systems.*

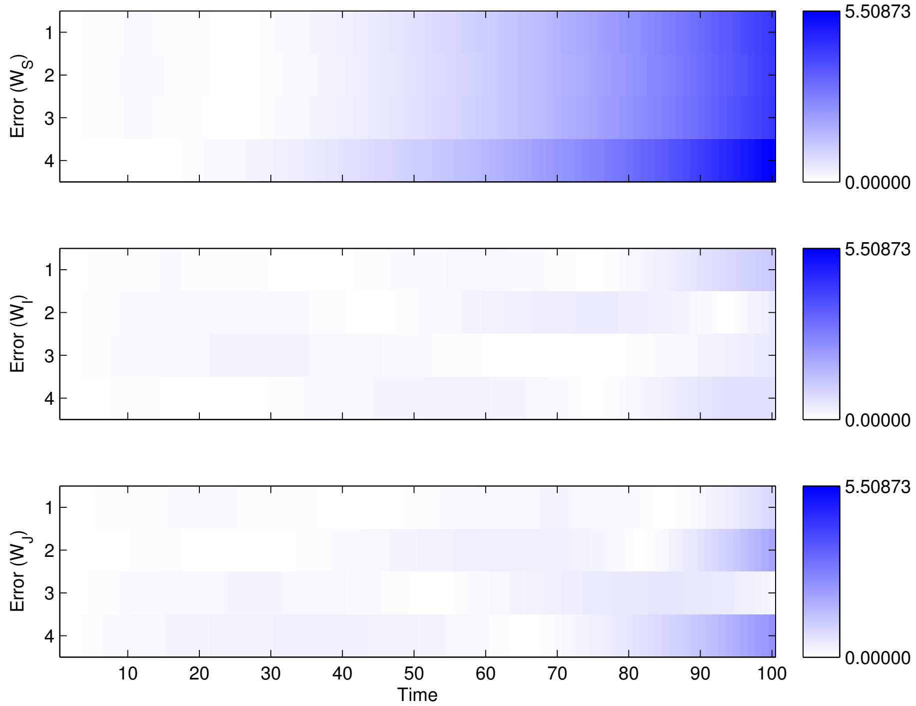

Figure 3: *Exemplary absolute error comparison of* ***parameter reduction*** *for a* ***linear*** *system by empirical sensitivity gramian, empirical identifiability gramian and empirical joint gramian.*

	offline time $[s]$	online time $[s]$	$L2$-error
original system	0	355.4448	0.0091
balanced truncation (W_S, W_O)	41.4663	12.1245	0.0423
balanced truncation (W_C, W_I)	136.3902	8.4331	0.0438
direct truncation (W_J)	1163.3232	8.0231	0.0360

Table 6: ***Nonlinear*** *Combined Reduction, averaged over 10 systems.*

Table 5 and table 6 show, along the the line of state and parameter reduction, the joint gramian fitting the original model closest. For either, the linear and nonlinear combined reduction, the joint gramian demands the most offline time while the online time is comparable to the balanced truncation methods.

8 Conclusion

In the case of nonlinear systems, the empirical cross gramian is advantageous compared to other gramian based model reduction methods. The cross gramian approximates, on average, in state, parameter and combined reduction the surrogate models with the least error. Moreover, the cross gramian can be applied to non-symmetric square systems with broader error bounds ([22, Remark

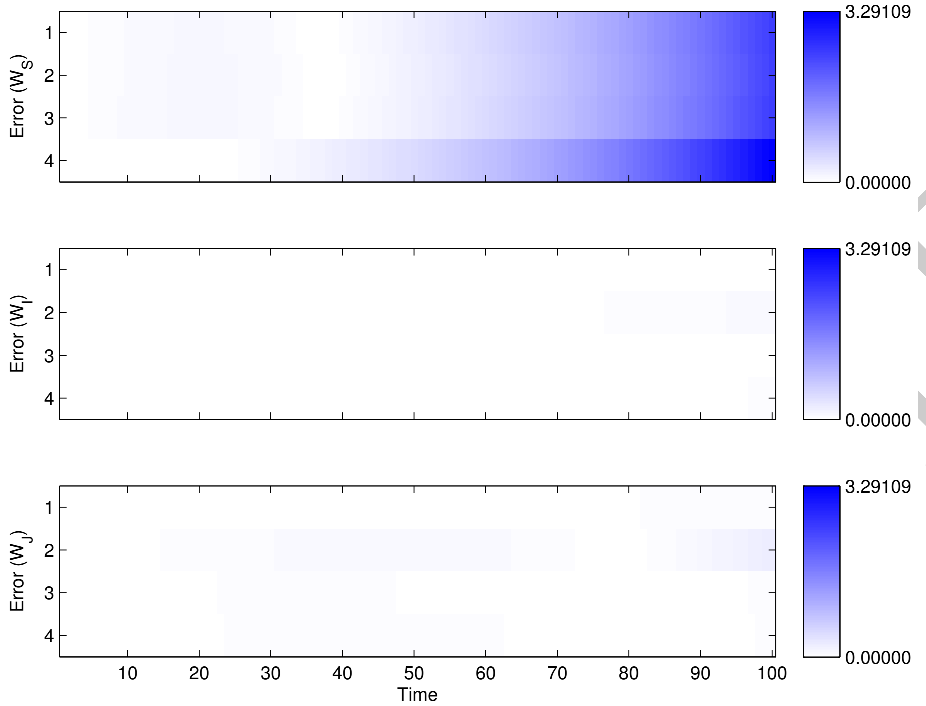

Figure 4: *Exemplary absolute error comparison of* ***parameter reduction*** *for a* ***nonlinear*** *system by empirical sensitivity gramian, empirical identifiability gramian and empirical joint gramian.*

2.1]). Surrogate models based on the joint gramian (parameter or combined reduction), though the computationally most costly, provide the best results in combined (state and parameter) reduction. Due to the longer assembly times of the joint gramian it is suggested to use it only if the dimension of the parameter space is of the same scale as the dimension of the state space. Further research has to be conducted on applying the empirical cross gramian to nonlinear or non-symmetric systems. Either by heuristically applying it ([23]), extending the symmetry constraint to nonlinear systems ([40]), or symmetrization ([21]).
Combined reduction of states and parameters using empirical gramians with either sensitivity and observability gramian, controllability and identifiability gramian or the joint gramian is applicable to (Bayesian) inverse problems. While the combination of sensitivity and observability gramian has the fastest offline time, the combination of controllability and identifiability gramian is more accurate and has lower online times. The joint gramian supersedes in accuracy which is bought by a high offline time.
The empirical cross gramian for MIMO systems completes the set empirical gramians for state reduction, while the joint gramian completes the body of parameter identification gramians.

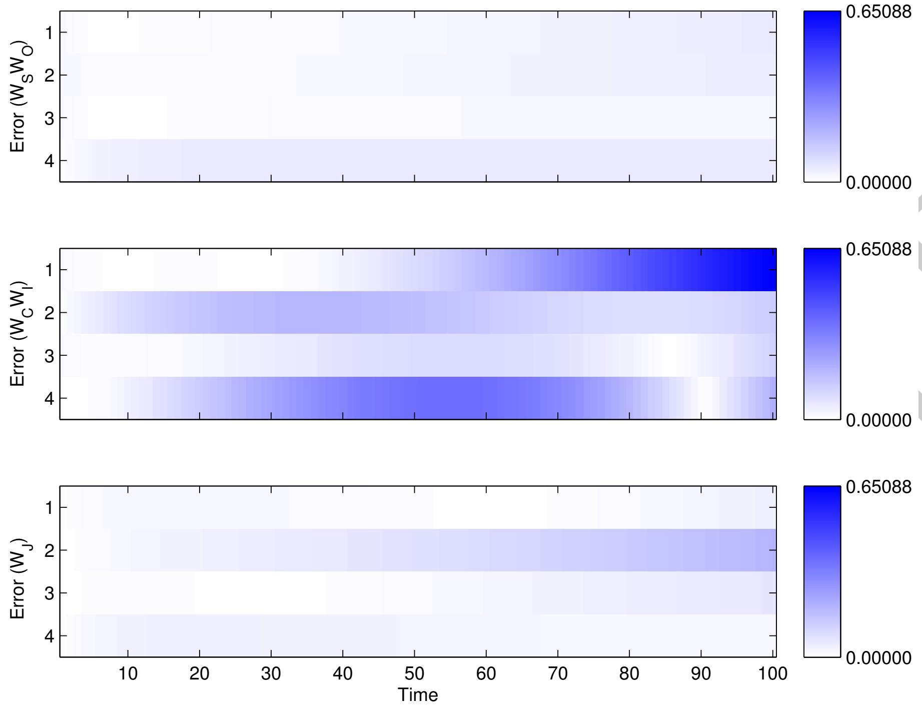

Figure 5: *Exemplary absolute error comparison of* ***combined state and parameter reduction*** *for a* ***linear*** *system by balanced truncation (with empirical sensitivity gramian and empirical observability gramian), balanced truncation (with empirical identifiability gramian and empirical controllability gramian) and empirical joint gramian.*

Acknowledgement

This work was supported by the Center Developing Mathematics in Interaction (DEMAIN) of the University of Muenster.

References

[1] P. M. Van Dooren. Gramian based model reduction of large-scale dynamical systems. *Numerical analysis 1999 (Dundee)*, 420:231–247, 2000.

[2] B. Moore. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE Transactions on Automatic Control*, 26(1):17–32, 1981.

[3] A. Laub, M. Heath, C. Paige, and R. Ward. Computation of system balancing transformations and other applications of simultaneous diagonalization algorithms. *IEEE Transactions on Automatic Control*, 32(2):115–122, 1987.

[4] A. Keil and J.L. Gouzé. Model reduction of modular systems using balancing methods. *Technical Report*, 2003.

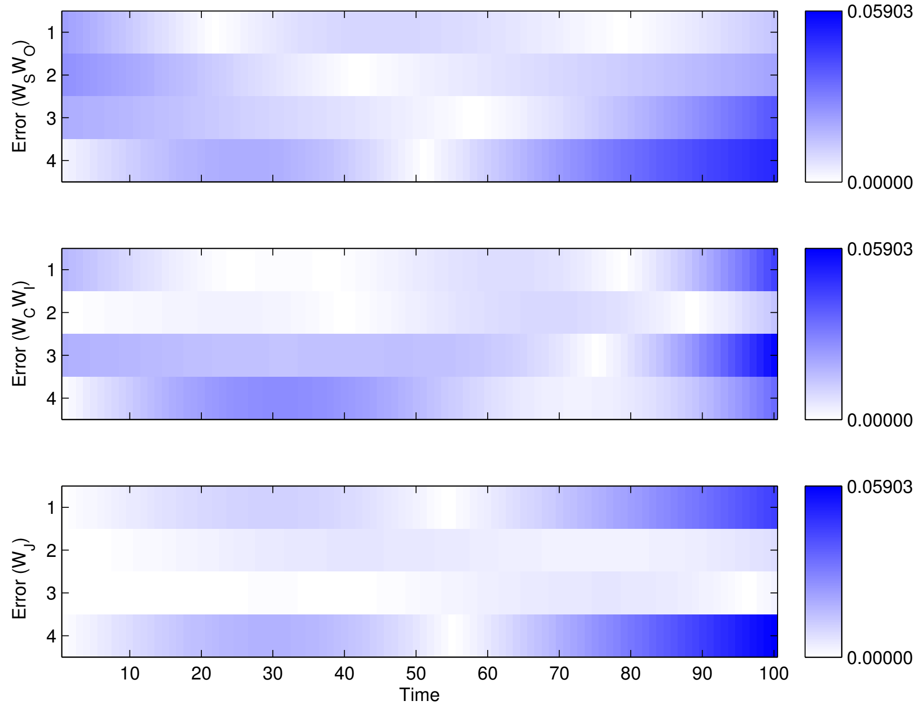

Figure 6: *Exemplary absolute error comparison of* ***combined state and parameter reduction*** *for a* ***nonlinear*** *system by balanced truncation (with empirical sensitivity gramian and empirical observability gramian), balanced truncation (with empirical identifiability gramian and empirical controllability gramian) and empirical joint gramian.*

[5] K.V. Fernando and H. Nicholson. On the structure of balanced and other principal representations of SISO systems. *Automatic Control, IEEE Transactions on*, 28(2):228 – 231, 1983.

[6] A.C. Antoulas. *Approximation of large-scale dynamical systems*, volume 6. Society for Industrial Mathematics, 2005.

[7] S. Lall, J.E. Marsden, and S. Glavaski. Empirical model reduction of controlled nonlinear systems. *Proceedings of the IFAC World Congress*, F:473–478, 1999.

[8] S. Lall, J.E. Marsden, and S. Glavaski. A subspace approach to balanced truncation for model reduction of nonlinear control systems. *International Journal of Robust and Nonlinear Control*, 12(6):519–535, 2002.

[9] J. Hahn and T.F. Edgar. A gramian based approach to nonlinearity quantification and model classification. *Industrial & engineering chemistry research*, 40(24):5724–5731, 2001.

[10] J. Hahn and T.F. Edgar. Balancing approach to minimal realization and model reduction of stable nonlinear systems. *Industrial & engineering chemistry research*, 41(9):2204–2212, 2002.

[11] S. Streif, R. Findeisen, and E. Bullinger. Relating cross Gramians and sensitivity analysis in systems biology. *Theory of Networks and Systems*, 10.4:437–442, 2006.

[12] S. Streif, S. Waldherr, F. Allgöwer, and R. Findeisen. Steady state sensitivity analysis of biochemical reaction networks: a brief review and new methods. In Arul Jayaraman and Juergen Hahn, editors, *Methods in Bioengineering*, pages 129–148. Artech House MIT Press, 2009. Systems Analysis of Biological Networks.

[13] D. Geffen, R. Findeisen, M. Schliemann, F. Allgower, and M. Guay. Observability based parameter identifiability for biochemical reaction networks. *Proceedings of the American Control Conference*, pages 2130–2135, 2008.

[14] C. Sun and J. Hahn. Parameter reduction for stable dynamical systems based on Hankel singular values and sensitivity analysis. *Chemical engineering science*, 61(16):5393–5403, 2006.

[15] K.V. Fernando and H. Nicholson. Minimality of SISO linear systems. *Proceedings of the IEEE*, 70(10):1241 – 1242, 1982.

[16] K. V. Fernando and H. Nicholson. On the Cauchy index of linear systems. *IEEE Trans. Automat. Control*, 28(2):222–224, 1983.

[17] K.V. Fernando and H. Nicholson. On a fundamental property of the cross-Gramian matrix. *Circuits and Systems, IEEE Transactions on*, 31(5):504–505, 1984.

[18] K.V. Fernando and H. Nicholson. Reachability, observability, and minimality of MIMO systems. *Proceedings of the IEEE*, 72(12):1820 – 1821, 1984.

[19] K.V. Fernando and H. Nicholson. On the cross-Gramian for symmetric MIMO systems. *Circuits and Systems, IEEE Transactions on*, 32(5):487–489, 1985.

[20] A.J. Laub, L.M. Silverman, and M. Verma. A note on cross-Grammians for symmetric realizations. *Proceedings of the IEEE*, 71(7):904–905, 1983.

[21] D.C. Sorensen and A.C. Antoulas. Projection methods for balanced model reduction. *Technical Report*, pages 1–18, 2001.

[22] D.C. Sorensen and A.C. Antoulas. The Sylvester equation and approximate balanced reduction. *Linear Algebra and its Applications*, 351:671–700, 2002.

[23] U. Baur and P. Benner. Cross-gramian based model reduction for data-sparse systems. *Electronic Transactions on Numerical Analysis*, 31:256–270, 2008.

[24] R. W. Aldhaheri. Model order reduction via real Schur-form decomposition. *Internat. J. Control*, 53(3):709–716, 1991.

[25] L. Fortuna and M. Frasca. *Optimal and Robust Control: Advanced Topics With MATLAB*. Taylor & Francis, 2012.

[26] J. Hahn and T.F. Edgar. Reduction of nonlinear models using balancing of empirical gramians and Galerkin projections. *Proceedings of the American Control Conference*, 4:2864 –2868, 2000.

[27] A.C. Antoulas and D.C. Sorensen. Approximation of large-scale dynamical systems: an overview. *Int. J. Appl. Math. Comput. Sci.*, 11(5):1093–1121, 2001.

[28] J. Hahn and T.F. Edgar. An improved method for nonlinear model reduction using balancing of empirical gramians. *Computers & chemical engineering*, 26(10):1379–1397, 2002.

[29] J. Hahn, T.F. Edgar, and M. Marquardt. Controllability and observability covariance matrices for the analysis and order reduction of stable nonlinear systems. *Journal of Process Control*, 13(2):115 – 127, 2003.

[30] M. Condon and R. Ivanov. Empirical balanced truncation of nonlinear systems. *Journal of Nonlinear Science*, 14(5):405–414, 2004.

[31] H. R. Shaker and M. Komareji. Control Configuration Selection for Multivariable Nonlinear Systems. *Industrial and Engineering Chemistry Research*, 51(25):8583–8587, 2012.

[32] A.K. Singh and J. Hahn. On the use of Empirical Gramians for Controllability and Observability Analysis. *Proceedings of the American Control Conference*, 2005.

[33] C. Himpe and M. Ohlberger. A Unified Software Framework for Empirical Gramians. *arXiv(math.OC)*, 1301.6879, 2013. (Preprint).

[34] C. Sun and J. Hahn. Model reduction in the presence of uncertainty in model parameters. *Journal of Process Control*, 16(6):645 – 649, 2006.

[35] A. Krishnamoorthy and K. Kocagoez. Singular Values using Cholesky Decomposition. *arXiv:cs.MS*, 2012.

[36] C. Eberle and C. Ament. Identifiability and online estimation of diagnostic parameters with in the glucose insulin homeostasis. *Biosystems*, 107(3):135 – 141, 2012.

[37] Y. Quan, H. Zhang, and L. Cai. Modeling and control based on a new neural network model. *Proceedings of the American Control Conference*, 3:1928–1929, 2001.

[38] K. Willcox and J. Peraire. Balanced model reduction via the proper orthogonal decomposition. *AIAA journal*, 40(11):2323–2330, 2002.

[39] P.J. Schmid and F. Gallaire. Model Reduction Techniques for Flow Control Applications. *2nd Young ERCOFTAC Workshop*, 2008.

[40] T. C. Ionescu, K. Fujimoto, and J. M. A. Scherpen. Singular value analysis of nonlinear symmetric systems. *IEEE Trans. Automat. Control*, 56(9):2073–2086, 2011.